

ADAPTIVE MEDIAN FILTERING FOR PREPROCESSING
OF TIME SERIES MEASUREMENTS

Matti Paunonen

Finnish Geodetic Institute
Ilmalankatu 1A
SF-00240 Helsinki, Finland
Telefax:358-0-264995

ABSTRACT.

A median (L1-norm) filtering program using polynomials was developed. This program was used in automatic recycling data screening. Additionally a special adaptive program to work with asymmetric distributions was developed. Examples of adaptive median filtering of satellite laser range observations and TV satellite time measurements are given. The program proved out versatile and time saving in data screening of time series measurements.

1. INTRODUCTION

The advantages of data screening of satellite laser range measurements using median (or L1-norm) instead of least squares were shown earlier (Paunonen 1989). The median is known to be insensitive to outlying observations, which is a useful property in preliminary data screening of any time series measurements. Asymmetric distributions often arise for various reasons; saturation of the laser detector and the receiver electronics, laser prepulses etc. This cannot be easily treated with least squares methods. In response to multiple needs, a median program permitting use of a higher order polynomial of up to ten was developed. A second version used automatic recycling of the fitting loop until a specified fit was obtained and a special program to work with unsymmetric distributions was devised. Examples of screening satellite laser range observations and TV satellite time measurements are given.

2.1 MEDIAN PROGRAM

The median program was constructed with the Fortran-procedure published by Barrodale and Roberts (1974), modified for polynomial use. The function to be minimized in the overdetermined case is

$$\sum_{i=1}^M |y_i - a_0 - a_1 x_i - a_2 x_i^2 - \dots - a_N x_i^N|, \quad (1)$$

where y_i are the observations, x_i the observing times (here), M the number of observations and N the degree of the polynomial to be fitted to the observations, and a_j ($j=0\dots N$) the coefficients of the polynomial.

2.2 AUTOMATIC MEDIAN FILTER PROGRAM

A versatile data screening program should be able to run automatically, without any manual interface. The median is good basis for a filtering program, because it selects reliably the densest part of the measurements as a reference. This means that any erratic points, that is outliers, are of true size, and not evened out as in the least squares method. Operation of the sequential median filtering program is started by forming the residuals of all the observations and calculating their average. Observations below a certain rejection limit are selected for the next cycle only. The problem is to find a suitable limit which is neither too inclusive nor too exclusive. Good operation was obtained by using the rejection level, R ,

$$R = 3.5 * AVR, \quad (2)$$

where AVR is the average of the residuals in the earlier round. For a Gaussian-shaped distribution the width between the zero and the point corresponding to one standard deviation is 1.46 times the width corresponding to the average. Thus the limit used is roughly equivalent to 2.4 times the standard deviation used in the least squares method. The repeated rounds are limited to four, but the final selection is generally ready after three rounds.

2.3 ADAPTIVE AUTOMATIC MEDIAN FILTER PROGRAM

In practice, the distribution of the data may be asymmetric and may include separated peaks. In satellite laser ranging, a distribution as shown in Fig.1 can easily arise with mode-locked lasers. The transmitted pulse may contain a small prepulse if selection of a single laser pulse from a train of mode-locked laser pulses is incomplete. Even if the parasitic pulse is small, it causes stops in the photon counting mode. This poses difficulties for normal screening methods using least squares. Use of only

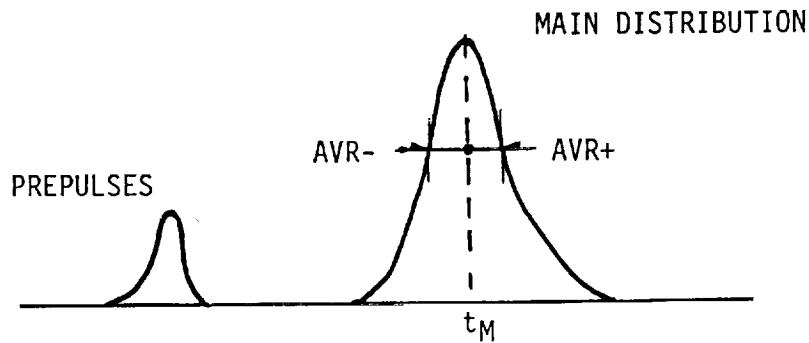


Fig. 1. An asymmetric distribution of possible range residuals in satellite laser ranging.

a two sigma-limit (Appleby and Sinclair 1991) is possible, but it may also fail if the contamination is severe. The median is expected to perform better because it finds the location of the main pulse more easily. The main part may also show unbalanced distribution (skewness) due, for instance, to saturation effects in the receiving electronics.

A modified method for asymmetric distributions is proposed as a refinement of the median filter described. Because the median produces separate average values for positive and negative residuals, AVR+ and AVR-, respectively, the program is allowed to select a minimum of the absolute values, and use it as the basis for the rejection limit in Eq. (2),

$$AVR = \text{Min}(|AVR+|, |AVR-|) . \quad (3)$$

At least mild skewness will be corrected in this way. If the distribution is symmetric, operation is normal.

3. TESTS OF THE ADAPTIVE MEDIAN FILTERING

The first test set was obtained from satellite laser range observations to the distant LAGEOS- satellite at Metsähovi, Fig. 2a. This is a mixture of good and bad observations. The points on the shorter range side arose from the shape of the laser pulse (Paunonen 1989). The short 4.5 ns pulse was cut by an electro-optical shutter from a 20 ns long ruby laser pulse. However, the shutter operation was not perfect and sometimes some leakage due, for instance, to changing temperature, may have occurred. This leakage, which was less than 10 per cent, looks like a pedestal on which a short pulse is riding.

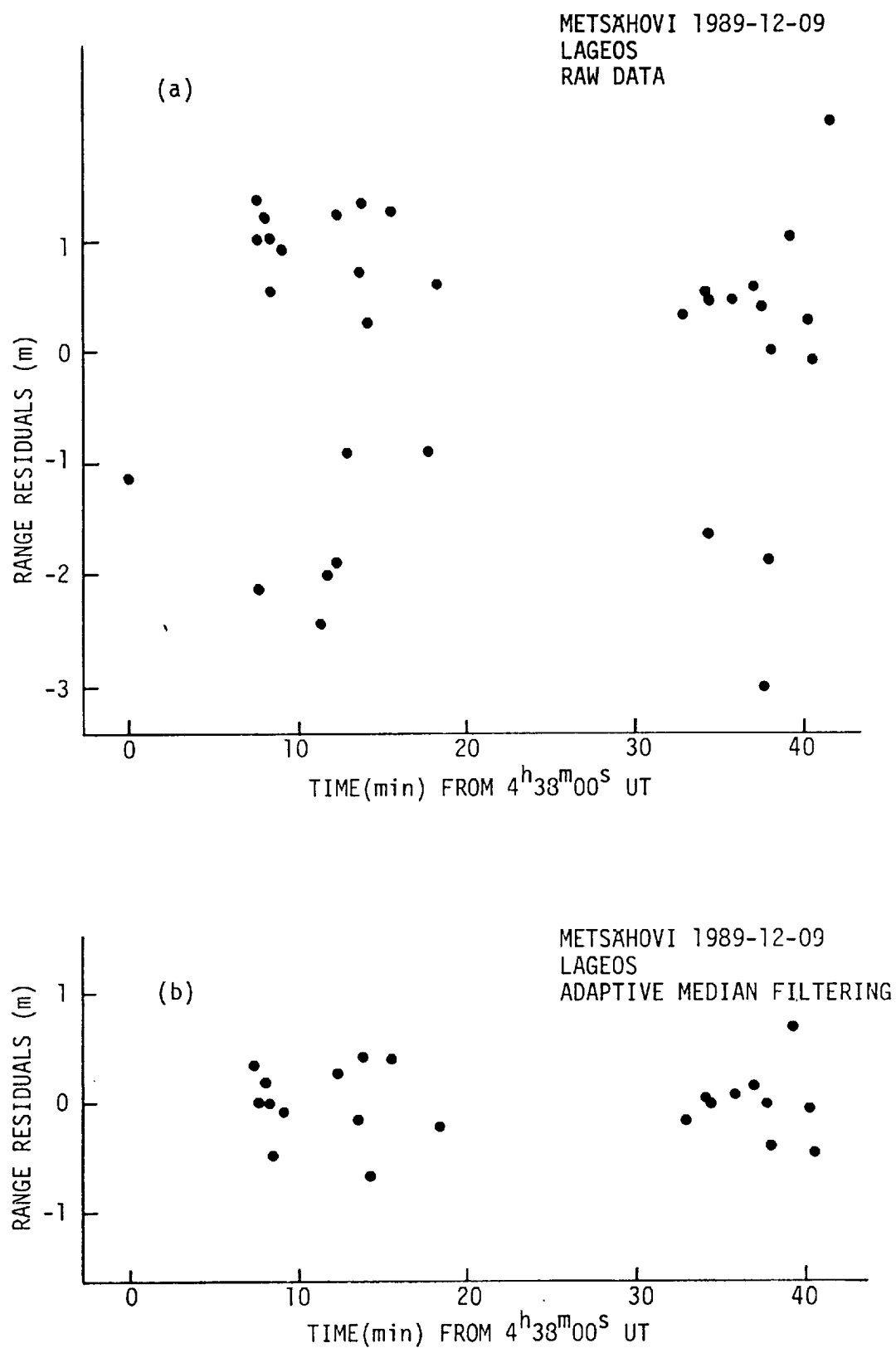


Fig. 2. a) Unscreened range residuals in a LAGEOS pass
b) Range residuals after adaptive median filtering
(linear fit).

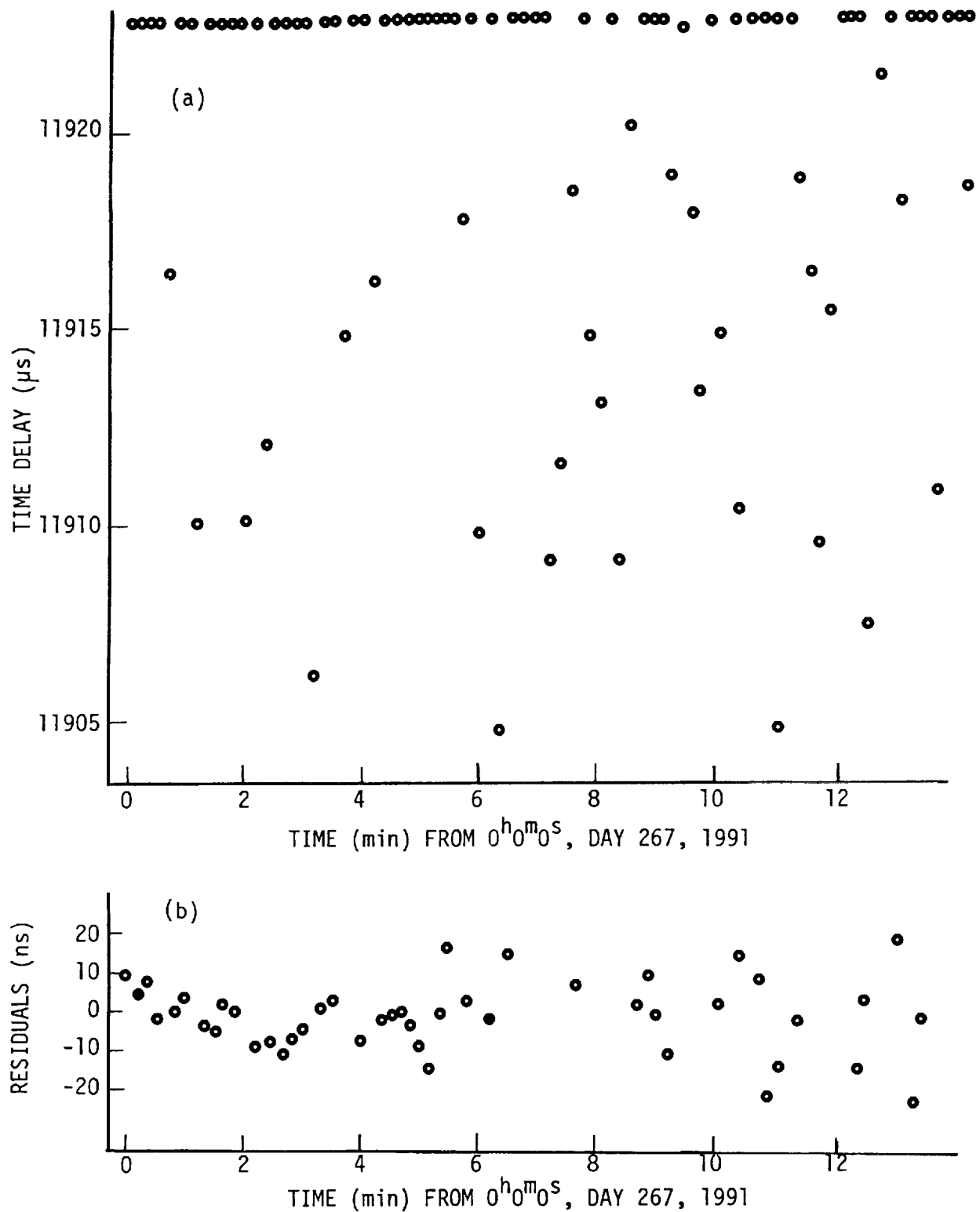


Fig. 3. a) Unscreened time measurements from a TV satellite
 b) Time residuals after adaptive median filtering (parabolic fit).

The energy ahead of the main pulse may stop the time interval counter. The adaptive median filter selected 22 observations from the 33 available ones (Fig. 2b). The initial data set is not random, because similar results were also obtained at the calibration line. The result is also plausible, because its width corresponds to the laser pulse width. The earlier method for testing one observation at a time (Paunonen 1982) would be of little value, without knowing what to seek.

Another test set originated from recent TV satellite time measurements (Fig. 3a). Here the second tick from the station clock started the time interval counter and the horizontal synchronization pulse from a satellite TV receiving system (50 Hz rate) stopped it. The large number of badly timed pulses probably originated from the encrypting method the TV transmission (program RAIUNO on the EUTELSAT 1-F5 satellite) is using. This set was also cleared well with the adaptive median filter (Fig. 3b). The r.m.s. value of the residuals was 11 ns.

4. DISCUSSION AND CONCLUSIONS

The adaptive median filter has proved to be versatile and to save time. It can safely remove several outliers, however large they are. This is a big advantage over the least squares method, in which all large outliers must be removed before any useful operations can be obtained. Polynomials should be used with care also in median filtering. End points in particular may behave peculiarly. The median program sets the value of some residuals at zero (this number is same as the degree of the polynomial), which is artificial. The asymptotic estimation efficiency of median is also usually worse than that of the mean (Eadie, et.al., 1971). It seems therefore best to use the median in the data screening phase and to use the normal least squares method for final extraction of the results. Use of the least squares is then well justified, because the distribution of the screened data is nearly normal.

REFERENCES

- Appleby, G.M. and A.T. Sinclair (1991). Formation of on-site normal points. Satellite Laser Ranging Newsletter, SLR Subcommittee of the CSTG (International Coordination of Space Techniques for Geodesy and Geodynamics), April 1991, pp. 25-28.
- Barrodale, I. and F.D.K. Roberts (1974). Solution of an over-determined system of equations in the l_1 norm. Communications of the ACM, Vol. 17, No. 6, pp. 319-320.

Eadie, W.T., D. Dryard, F.E. James, M. Roos, B. Sadoulet
(1971). Statistical Methods in Experimental Physics.
North-Holland Publishing Company, Amsterdam 1971,
296 pp.

Paunonen, M. (1982). Studies on the Metsähovi satellite
laser ranging system. Reports of the Finnish Geodetic
Institute, 1982:2, Helsinki 1982, 47 pp.

Paunonen, M. (1989). Use of the L1-norm in screening laser
range observations to satellites. Proc. of the 18th
INTERKOSMOS Symposium on the Use of Artificial Satellite
Observations for Geodesy and Geophysics, June 12-17,
1989, Cracow, Poland, (Ed. WL. Goral, Polish Academy of
Sciences, Observations of Artificial Satellites of the
Earth, No 27, Warsaw 1990), pp. 359-362.